Information for Other Resource Providers

Created: September 14, 2005
Updated: August 07, 2006

LinkOut is a feature of Entrez where third parties provide information to link specific Entrez records to relevant web-accessible online resources, such as full-text publications, molecular biology databases (i.e., organism-specific, taxonomy, structure, etc.), catalogs of research materials (clones, cell cultures, primers, etc.), funding sources, medical resources, research groups, and others. This document explains how providers of resources other than online full text can participate in LinkOut by supplying NCBI with the necessary information for creating links from Entrez records to the providers' resources.
An FAQ for LinkOut Providers is also available.

How It Works

LinkOut provides links from PubMed records and other Entrez database records to online resources external to the Entrez system. All linking information is submitted by LinkOut providers - the owner or agent for the owner of the online resource. LinkOut providers are responsible for maintaining their links.

To submit links to your resource, you will need to upload two XML files, an identity file and a resource file. The identity file contains the information about your organization needed to list your resource(s) in LinkOut. The resource file describes the Entrez records you will link from and contains the information that LinkOut needs to generate the links.

All links can be viewed in the LinkOut Display. In PubMed, links can also be viewed as icons.

Links submitted to LinkOut are available through the Entrez Programming Utility ELink. Thus, systems other than Entrez may retrieve links through this utility and display them at their sites.

Prerequisites for Participation

Resources submitted for inclusion in LinkOut will be evaluated individually to determine whether they meet certain inclusion criteria.

Resources eligible for linking from Entrez databases must be directly relevant to the specific subjects of the Entrez records and useful to users' study and research. Resources from professional societies, government agencies, educational institutions, or individuals and organizations that have received grants from major funding organizations are preferred.

Please review the Guidelines for Evaluation of Resources before applying for inclusion in LinkOut. Resources with a commercial interest should pay particular attention to the Additional Information for Commercial Interests section of the Guidelines.

Apply for Inclusion in LinkOut

To apply for inclusion in LinkOut, send an email to linkout@ncbi.nlm.nih.gov. Include the following information:

- Name, email address, and phone number of a contact person in your organization.

- The scope of your resource, including the URL of the resource. If a username and password are required to access the resource, please include a temporary username and password that the LinkOut team can use to evaluate the resource. Also, please describe the type of Entrez records to which you would like to apply links.

- Describe any restrictions on access to the resource.

A LinkOut team member will email the contact person within 1 week regarding your request.

File Preparation: Identity File

To create links to your resource, NCBI needs two files, an identity file and a resource file.

The identity file contains the information needed to list a provider in LinkOut. This file must be named **providerinfo.xml**; the file name is case sensitive. This file should be composed in a text editor, such as NotePad, not in a word processing program.

For complete file specifications and description of elements, see Technical Information: Identity File or the LinkOut DTD.

The following is an example providerinfo.xml file for the LinkOut participant, WebDatabase Co., with Provider Id **7777** and NameAbbr **WebDB**.

```
<?xml version="1.0"?>
<!DOCTYPE Provider PUBLIC "-//NLM//DTD LinkOut 1.0//EN"
"http://www.ncbi.nlm.nih.gov/entrez/linkout/doc/LinkOut.dtd">
<Provider>
<!-- ProviderId is assigned by NCBI -->
    <ProviderId>7777</ProviderId>
    <Name>WebDatabase Co.</Name>
    <NameAbbr>WebDB</NameAbbr>
    <SubjectType> gene/protein/disease-specific</SubjectType>
    <Attribute>registration required</Attribute>
<!-- Url is used in My NCBI and in the lists of LinkOut Providers -->
    <Url>http://www.webdatabase.com</Url>
<!-- Brief is used in My NCBI -->
    <Brief> On-line publisher of biomedical databases</Brief>
</Provider>
```

<SubjectType> and <Attribute> elements included in the providerinfo.xml file will apply to all links submitted by the provider. In the example above, access to all databases published by WebDatabase Co. requires a free registration, therefore <Attribute>registration required</Attribute> has been included in the providerinfo.xml file.

File Preparation: Resource File (XML)

Section Contents:

- Resource File Format

- <ObjectList>: Selecting Records in a Resource File

    - Selecting Records Using <ObjId>

    - Selecting Records Using <Query>

- Tips for Using <Query>

- Additional Information on Using <Query> When Linking to PubMed Citations

- <ObjectUrl>: Specifying the Link to Access the Resource

  - Creating the URL for the Link

  - Describing the Resource

- Resource File Examples

To create links to your resource, NCBI needs two XML files, an identity file and a resource file.

The resource file describes the Entrez records to which your links will be applied and contains the information that LinkOut needs to generate the links. Links described in the resource file must link directly to the relevant resource, requiring no additional searching to access the resource after a user clicks on the provider's link.

Resource files are typically submitted in XML format, or you may choose to submit links as a simple text file instead. Providers of online full text must submit resource files in XML format.

XML: Resource File Format

XML resource files must have a file extension **.xml**; the file extension is case sensitive. File names may contain alpha-numeric characters and underscores only. Special characters and spaces are not allowed. Typically, files are named **resources.xml**. To help with file management, a provider may supply more than one resource file. File size may not exceed 20 MB. This file should be composed in a text editor, such as NotePad, not in a word processing program.

For complete file specifications and description of elements, see LinkOut Technical Information: Resource File or the LinkOut DTD.

The resource file below describes links from NCBI's Nucleotide database to a *C. elegans* sequence database provided by WebDatabase Co., ProviderId 7777.

```
<?xml version="1.0"?>
<!DOCTYPE LinkSet PUBLIC "-//NLM//DTD LinkOut 1.0//EN"
"http://www.ncbi.nlm.nih.gov/entrez/linkout/doc/LinkOut.dtd"
[ <!ENTITY icon.url "http://www.webdatabase.com/images/webdb.gif">
<!ENTITY base.url "http://www.webdatabase.com/cgi-bin/elegans?">]>
<LinkSet>
 <Link>
   <LinkId>1</LinkId>
   <ProviderId>7777</ProviderId>
   <ObjectSelector>
     <Database>Nucleotide</Database>
     <ObjectList>
       <Query>Caenorhabditis elegans [orgn]</Query>
     </ObjectList>
   </ObjectSelector>
   <ObjectUrl>
     <Base>&base.url;</Base>
     <Rule>an_lookup=&lo.pacc;</Rule>
       <UrlName>Caenorhabditis elegans</UrlName>
       <SubjectType>organism-specific</SubjectType>
   </ObjectUrl>
 </Link>
</LinkSet>
```

<ObjectList>: Selecting Records in a Resource File

The <ObjectList> element is used to select the Entrez records to which links will be applied. <ObjectList> contains one or more <Query> elements OR one or more <ObjId> elements. <Query> elements contain a valid Entrez search that will retrieve the records to which the link described in <ObjectUrl> will be applied. <ObjId> elements contain the Unique Identifier of the Entrez records to which the link described in <ObjectUrl> will be applied.

Selecting Records Using <ObjId>

<ObjId> contains the unique identifier for a record in an Entrez database. For example, GI for Nucleotide and Protein.

Example: Select record with GI 6016240 in Entrez Protein

```
<ObjectList>
<ObjId>6016240</ObjId>
</ObjectList>
```

More than one <ObjId> can be used in an <ObjectList>.

Example: Select records with GIs 6016240 and 6016241

```
<ObjectList>
<ObjId>6016240</ObjId>
<ObjId>6016241</ObjId>
</ObjectList>
```

Because <ObjId> requires more maintenance than <Query>, NCBI recommends using <Query> whenever possible. When <ObjId> is used, the provider is responsible for updating the holdings file as new records are submitted to PubMed.

Selecting Records Using <Query>

The <Query> element contains a valid Entrez search. See Entrez Help for information on constructing Entrez queries. Links will be applied to the citations retrieved by the search.

Tips for Using <Query>

1. Ranging is not allowed in Unique Identifier searches. For journal searches, ranging is additionally not allowed in Volume, Issue, or Page searches.

2. Truncation is not allowed in search statements.

3. Search field tags must be enclosed in square brackets.

4. To include a date range in searches, use this format: startyear:endyear[dp]. Dates should be notated as YYYY/MM/DD. Month and Day are optional.

5. Do not use the search field tags [sb] or [filter].

6. Boolean operators AND, OR, NOT must be in uppercase.

7. Use either NLM's Title Abbreviations [ta] or ISSN numbers in journal searches. Title Abbreviations must be entered in double quotes, e.g., "J Mol Dis" [ta].

> Example: Select records with the organism "Caenorhabditis elegans" published from 1996 to 1999 in Entrez Nucleotide
> <Database>Nucleotide</Database><ObjectList>
> <Query>Caenorhabditis elegans [orgn] AND 1996:1999 [pdat]</Query>
> </ObjectList>

See the results of this <Query> in Entrez Nucleotide.

> Example: Select records with the organism "Caenorhabditis elegans" published by J. Smith in Entrez Nucleotide. As new records are
> submitted to Entrez, links will be applied automatically.
> <Database>Nucleotide</Database><ObjectList>
> <Query>Caenorhabditis elegans [orgn] AND smith j [auth]</Query>
> </ObjectList>

See the results of this <Query> in Entrez Nucleotide.

More than one <Query> can be listed within the <ObjectList>, as shown in the example below.

> Example: Select records for chimpanzees starting from the publication date January 1, 2000 and records for humans starting from
> January 1, 2002 in Entrez Genome. As new records are submitted, links will be applied automatically.
> <Database>Genome</Database>
> <ObjectList>
> <Query>chimpanzee [orgn] AND 2000:2010[dp]</Query><Query>human [orgn] AND 2002:2010[dp] </Query>
> </ObjectList>

See the results for this <ObjectList> in Entrez Genome.

Additional Information on Using <Query> When Linking to PubMed Citations

Links in PubMed may be created using MeSH headings. In this case, the <Query> should be very precise. Only Major headings should be used in the <Query> and noexp should be used so the terms will not be exploded. See the MeSH database for information about using MeSH and tutorials.

> Example: Use MeSH headings to select citations on acupuncture therapy in the treatment of tension headaches in PubMed.
> <Database>PubMed</Database><ObjectList>
> <Query>"Acupuncture Therapy/methods"[majr:noexp] AND "Tension Headache/therapy"[majr:noexp]</Query>
> </ObjectList>

See the results for this <Query> in PubMed.

<ObjectUrl>: Specifying the Link to Access the Resource

The <ObjectUrl> element is used to describe the link to the online resource. <ObjectUrl> contains the sub-elements <Base>, <Rule>, <SubjectType>, <Attribute>, and <UrlName>. <Base> and <Rule> are concatenated to form the URL for the link. <SubjectType>, <Attribute>, and <UrlName> describe the resource to which the record is being linked.

For definitions of SubjectType and Attribute elements, see Special Elements: SubjectType and Special Elements: Attribute.

Creating the URL for the Link

<Base> is the stable portion of the URL for the provider's resource. This is usually the URL of the provider's website or CGI program.

<Rule> is the remainder of the URL needed to access the appropriate record within the resource.

> Example: Create the URL http://www.webdatabase.com/cgi-bin/elegans?OID=1988
> <Base>http://www.webdatabase.com/ cgi-bin/elegans?<Base>
> <Rule>OID=1988</Rule>

If the URL for the resource follows a pattern using variable values that are found in the Entrez record, the pattern can be described in the <Rule> element, and LinkOut can insert the appropriate values for each citation. This allows many links to be generated from the information in a single <ObjectUrl>.

URL patterns are described using LinkOut's XML entities. An XML entity is a short text string that represents a type of value. During LinkOut processing, the text string is replaced in the URL by the appropriate value for each record. Supported entities can be found in the list of Allowable Rule Keywords and in the LinkOut DTD.

> Example: Create URLs following the pattern:
> http://www.webdatabase.com/cgi-bin/an_lookup=[PACC]
> <Base>http://www.webdatabase.com/cgi-bin/</Base>
> <Rule> an_lookup=&lo.pacc;</Rule>

Using this <Base> and <Rule>, the URL constructed for the record with accession number AL032671 would be *http://www.webdatabase.com/cgi-bin/elegans?an_lookup=AL032671*

Entities can be combined with other information in the <Rule>.

> Example: Create URLs following the pattern:
> http://www.webdatabase.com/cgi-bin/db=elegans&id_lookup=[Entrez Unique Identifier]&view=text
> <Base>http://www.webdatabase.com/cgi-bin</Base>
> <Rule>/db=elegans&id_lookup=&lo.id;&view=text</Rule>

In this case, the URL generated for the record with the unique ID "6016240" would be: http://www.webdatabase.com/cgi-bin/db=elegans&id_lookup=6016240&view=text

LinkOut does not support Unicode (UTF-8) and requires that certain special characters be encoded in files. See Special Characters for a list of these characters and their encodings.

Describing the Resource

The relevance of resources linked from Entrez records should be readily apparent to users. The name and/or description of the resource should convey something about the information that is being offered and its relevance.

The elements <SubjectType> and <Attribute> are used in the <ObjectUrl> to describe resources. Available SubjectTypes can be found in Special Elements: SubjectType. Available Attributes can be found in Special Elements: Attribute.

If the available SubjectTypes and Attributes do not suffice to describe the resource, UrlName can be used as well. If no SubjectType is given, the SubjectType "miscellaneous" will be assigned automatically.

The application of SubjectTypes and Attributes is at the discretion of the resource provider. However, if there are any barriers to accessing the resource, one of the following Barrier Attributes must be used:

```
<Attribute>registration required</Attribute>
<Attribute>subscription/membership/fee required</Attribute>
```

Continuing the example above, if WebDatabase Co. requires a subscription to access the *C. elegans* database, the <ObjectUrl> element might look like this:

```
<ObjectUrl>
<Base>http://www.webdatabase.com/cgi-bin/</Base>
<Rule>/db=elegans&amp;id_lookup=&lo.id;&amp;view=text</Rule>
<Attribute>subscription/membership/fee required</Attribute>
</ObjectUrl>
```

Resource File Examples

Example 1: Molecular Biology Database, Inc., Provider Id 1234, provides links to freely available information for the Taxonomy records with IDs 9606 and 111063. URLs for the database entries are created using a text string that is not included in Taxonomy, so links are created individually for each record. To minimize the repetition of textual data, the Base for the URL has been defined as an Entity in the Prolog of the file.

```
<!DOCTYPE LinkSet PUBLIC "-//NLM//DTD LinkOut 1.0//EN"
"http://www.ncbi.nlm.nih.gov/entrez/linkout/doc/LinkOut.dtd"
[<!ENTITY base.url "http://molbioco.com/animals/">]>
<LinkSet>
 <Link>
   <LinkId>1</LinkId>
   <ProviderId>1234</ProviderId>
   <ObjectSelector>
      <Database>taxonomy</Database>
      <ObjectList>
         <ObjId>9606</ObjId>
      </ObjectList>
   </ObjectSelector>
   <ObjectUrl>
      <Base>&base.url;</Base>
      <Rule>homo/h._sapiens</Rule>
      <UrlName>Homo sapiens</UrlName>
      <SubjectType>taxonomy/phylogenetic</SubjectType>
   </ObjectUrl>
 </Link>

 <Link>
   <LinkId>2</LinkId>
   <ProviderId>1234</ProviderId>
   <ObjectSelector>
      <Database>taxonomy</Database>
      <ObjectList>
         <ObjId>9733</ObjId>
      </ObjectList>
   </ObjectSelector>
   <ObjectUrl>
      <Base>&base.url;</Base>
      <Rule>orcinus/o._orca</Rule>
      <UrlName>Orcinus orca</UrlName>
      <SubjectType>taxonomy/phylogenetic</SubjectType>
    </ObjectUrl>
 </Link>
</LinkSet>
```

Example 2: Genotypes, Inc., Provider Id 4321, provides free online access to genotyping assays from records in the Entrez database SNP. SNP records are selected using the SNP unique identifier. The URL to access the assays at their site follows this pattern for each record: http://gti.com/Gateway?source= SNP&res=Assays&ap1=rs[SNP ID]

```
<?xml version="1.0"?>
<!DOCTYPE LinkSet PUBLIC "-//NLM//DTD LinkOut 1.0//EN"
"http://www.ncbi.nlm.nih.gov/entrez/linkout/doc/LinkOut.dtd"
[<!ENTITY base.url "http://gti.com/Gateway?source=SNP&amp;res=Assays&amp;">]>
<LinkSet>
 <Link>
    <LinkId>1</LinkId>
    <ProviderId>4321</ProviderId>
    <ObjectSelector>
       <Database>SNP</Database>
       <ObjectList>
          <ObjId>7928656</ObjId>
          <ObjId>2049045</ObjId>
          <ObjId>1811350</ObjId>
          <ObjId>1871598</ObjId>
          <ObjId>7947824</ObjId>
          <ObjId>681267</ObjId>
          <ObjId>1947741</ObjId>
       </ObjectList>
    </ObjectSelector>
    <ObjectUrl>
       <Base>&base.url;</Base>
       <Rule>ap1=rs&lo.id;</Rule>
       <UrlName>Genotyping Assays</UrlName>
    </ObjectUrl>
 </Link>
</LinkSet>
```

Example 3: A record may be retrieved by more than one <Query>. When this happens, link assignment will be handled as described in Duplicate Links and Multiple Links.

If these queries are in different Link elements, <Attribute>preference</Attribute> can be used to indicate which <Link> element should be applied to the record. This is generally used in situations where the links for a subset of a range have a different URL pattern or different access restrictions. See Duplicate Links and Multiple Links for more information.

In the example below, the records included in LinkId 1 will also be selected by LinkId 2.

The LinkOut provider WebDatabase Co. provides links from Entrez Nucleotide to the *C. elegans* sequence database.

LinkId 1 describes links from particular Nucleotide records, specified by unique ID (GI) to a single record in the *C. elegans* database. The records are selected using <ObjId> and have a special <Rule>. Because these records are also included in LinkId 3, <Attribute>preference</Attribute> is used to indicate that only this link should be applied to these citations.

LinkId 2 provides links from all Nucleotide records on *C. elegans* to WebDatabase Co.'s *C. elegans* records, except for the records selected in LinkId 1 and LinkId 2.

As **LinkId 1** describes specific requirements, it is listed before the general **LinkId 2**.

```
<?xml version="1.0"?>
<!DOCTYPE LinkSet PUBLIC "-//NLM//DTD LinkOut 1.0//EN"
"http://www.ncbi.nlm.nih.gov/entrez/linkout/doc/LinkOut.dtd"
[<!ENTITY base.url "http://www.webdatabase.com/cgi-bin/elegans?">]>
<LinkSet>
 <Link>
    <LinkId>1</LinkId>
```

```
     <ProviderId>7777</ProviderId>
     <ObjectSelector>
       <Database>Nucleotide</Database>
       <ObjectList>
         <ObjId>3810674</ObjId>
         <ObjId>1217583</ObjId>
         <ObjId>1181594</ObjId>
       </ObjectList>
     </ObjectSelector>
     <ObjectUrl>
       <Base>&base.url;</Base>
       <Rule>db=special&amp;ID=A594E</Rule>
       <Attribute>preference</Attribute>
     </ObjectUrl>
   </Link>

   <Link>
     <LinkId>3</LinkId>
     <ProviderId>7777</ProviderId>
     <ObjectSelector>
       <Database>Nucleotide</Database>
       <ObjectList>
         <Query>Caenorhabditis elegans [orgn]</Query>
       </ObjectList>
     </ObjectSelector>
     <ObjectUrl>
       <Base>&base.url;</Base>
       <Rule>an_lookup=&lo.pacc;&amp;view=full</Rule>
     </ObjectUrl>
   </Link>
   </LinkSet>
```

File Preparation: Resource File (Simple Text)

Section Contents:

- Resource File Format

- Global Information

- Creating Links

    - Selecting Records

    - Specifying the Link

    - Describing the Resource

- Resource File Examples

Simple Text: Resource File Format

Providers may choose to submit the resource file in a simple text file instead of XML, for easier processing.

Text resource files must have a file extension **.ft**; the file extension is case sensitive. File names may contain alpha-numeric characters and underscores only. Special characters and spaces are not allowed. Typically, files are named "resources.ft". To help with file management, a provider may supply more than one resource file. File size may not exceed 10 MB.

Text files will be translated into valid XML files that conform to the LinkOut DTD and then are processed into LinkOut. This file should be composed in a text editor, such as NotePad, not in a word processing program.

For complete file specifications and description of elements, see Resource File: Simple Text or the LinkOut DTD.

The resource file below describes links from NCBI's Nucleotide database to a *C. elegans* sequence database provided by WebDatabase Co., ProviderId 7777.

```
--- lines starting with "-" are comments ---
--- information in the first block is global ---
prid:   7777
dbase:  nucleotide
stype:  organism-specific
!base:  http://www.webdatabase.com/cgi-bin/elegans?

------
linkid: 1
query:  Caenorhabditis elegans [orgn]
base:   &base;
rule:   an_lookup=&lo.pacc;
name:   Caenorhabditis elegans
------
```

Simple Text: Global Information

The first block holds global information that will be used throughout the file.

prid: LinkOut Provider ID

dbase: Entrez database that will be hosting the links, e.g., pubmed, nucleotide, taxonomy

stype: SubjectType. See Special Elements: SubjectType for all available SubjectTypes.

attr: Attribute. See Special Elements: Attribute for all available Attributes.

Simple Text: Creating Links

Each subsequent block specifies a LinkOut link. This has two basic parts, an Entrez query and a URL pointing back to the provider's site. Each Entrez query will be evaluated, and a link to the specified URL will be applied to records that are retrieved by the query. See Entrez Help for information on constructing Entrez queries.

In the simplest case, each block could be:

```
-----
query: [a valid Entrez query]
rule:  [the URL that will be applied to the records retrieved by the query]
-----
```

Simple Text: Selecting Records

Any valid Entrez query may be used to select records. See Entrez Help for information on constructing Entrez queries. Each query should appear on a single **query** line. Multiple query lines in one block will be OR-ed together:

Example: This search will be translated as: human[name] OR chimpanzee[name]
query: human [name]
query: chimpanzee [name]

Each record in an Entrez database has a numerical unique identifier (UID). You can select the Entrez records that you would like to link from by UID in the **uids:** line.

> Example: Place links on records with UIDs 123456, 123469, and 3847559
> ---- separate unique identifiers (UIDs) with a blank space. Each new line should start with the "uids:" label ---
> uids: 123456 123469 3847559

Simple Text: Specifying the Link

The link is specified using the **base:** and **rule:** lines. **base:** is the stable portion of the URL for the resource. This is usually the URL of the provider's website or CGI program. **rule:** is the remainder of the URL needed to access the online resource.

> **base:** and **rule:** are concatenated to form the URL for the link.

> If desired, the entire URL for the resource can be included on the **rule:** line, and the **base:** line can be omitted.

> Example: The following will both create a link to the URL http://www.webdatabase.com/cgi-bin/elegans?OID=1988
> ------
> rule: http://www.webdatabase.com/ cgi-bin/elegans?OID=1988
> -------
> ------
> base: http://www.webdatabase.com/cgi-bin/elegans?
> rule: OID=1988
> ------

> If the URL for the resource follows a pattern using variable values that are found in the Entrez record, the pattern can be described on the **rule:** line, and LinkOut can insert the appropriate values for each citation. This allows many links to be generated from the information in a single block.

> URL patterns are described using LinkOut's XML entities. An XML entity is a short text string that represents a type of value. During LinkOut processing, the text string is replaced in the URL by the appropriate value for each record. Supported entities can be found in the list of LinkOut Entities and in the LinkOut DTD.

> Example: Create URLs following the pattern: http://www.webdatabase.com/cgi-bin/an_lookup=[PACC]
> **base:** http://www.webdatabase.com/cgi-bin/
> **rule:** an_lookup=&lo.pacc;

> Using this **base:** and **rule:**, the URL constructed for the record with accession number AL032671 would be *http://www.webdatabase.com/cgi-bin/elegans?an_lookup=AL032671*

> Entities can be combined with other information in the **rule:**

> Example: Create URLs following the pattern:
> http://www.webdatabase.com/cgi-bin/db=elegans&id_lookup=[Entrez Unique Identifier]&view=text
> **base:** http://www.webdatabase.com/cgi-bin
> **rule:** db=elegans&id_lookup=&lo.id;&view=text

In this case, the URL generated for the record with the unique ID 6016240 would be: *http://www.webdatabase.com/cgi-bin/db=elegans&id_lookup=6016240&view=text*

To minimize the repetition of textual data, the **base:** portion of the URL can be defined as an entity in the global information block, as shown below.

```
prid: 4592
dbase: PubMed
!base.url: http://a257.g.akamaitech.net/7/257/2422/
------------------------------------------------------------------
linkid: 704411419
uids: 15754467
base: &base.url;
rule: 01jan20051800/edocket.access.gpo.gov/2005/pdf/05-4062.pdf
attr: full-text PDF
------------------------------------------------------------------
linkid: 70389516
uids: 15736310
base: &base.url;
rule: 01jan20051800/edocket.access.gpo.gov/2005/pdf/05-3829.pdf
attr: full-text PDF
------------------------------------------------------------------
linkid: 70379232
uids: 15732197
base: &base.url;
rule: 01jan20051800/edocket.access.gpo.gov/2005/pdf/05-3728.pdf
attr: full-text PDF
------------------------------------------------------------------
```

Simple Text: Describing the Resource

The relevance of resources linked from Entrez records should be readily apparent to users. The name and/or description of the resource should convey something about the information that is being offered and its relevance.

The following optional fields allow you to describe your links and resources.

icon: URL of an icon file that you would like to represent your link and resources. Only applicable to links in PubMed. The icon should meet the specifications described in Icons.

name: Additional description of the link. **name:** should only be used when the values in the LinkOut SubjectType and Attribute lists are not sufficient.

stype: SubjectType. See Special Elements: SubjectType for all available SubjectTypes. SubjectType is used to determine where links will be placed in the LinkOut Display. If no SubjectType is given, the SubjectType "miscellaneous" will be assigned automatically.

attr: Attribute. See Special Elements: Attribute for all available Attributes.

The application of SubjectTypes and Attributes is at the discretion of the resource provider. However, if there are any barriers to accessing the resource, one of the following Barrier Attributes*must* be used:

registration required
subscription/membership/fee required

Simple Text: Resource File Examples

Example 1: The following file shows five links to taxonomic resource on the Web. These links can be seen in Entrez Taxonomy. The links are displayed under "NCBI taxonomy bookmarks" in the LinkOut Display.

Because each link has an individual URL, the links are made separately.

```
----------NCBI taxonomy bookmark links ----
prid:   3206
dbase:  Taxonomy
------Apis mellifera)------
linkid: 1
```

```
query:  Apis mellifera [name]
rule:   http://beelab.cas.psu.edu/intro.html
name:   Honey Bee Lab (Penn State)
------
linkid: 2
query:  Apis mellifera [name]
rule:   http://www.barc.usda.gov/psi/brl/
name:   Bee Research Lab (USDA Beltsville)
------
linkid: 3
query:  Apis mellifera [name]
rule:   http://www.hgsc.bcm.tmc.edu/projects/honeybee/
name:   Baylor Honey Bee Genome
------
linkid: 4
query:  Apis mellifera [name]
rule:   http://titan.biotec.uiuc.edu/bee/honeybee_project.htm
name:   Honey Bee Brain EST Project
------
linkid: 5
query:  Apis mellifera [name]
rule:   http://ourworld.compuserve.com/homepages/Beekeeping/weblinks.htm
name:   Bee Web Links
------
```

Example 2: The hypothetical provider Genotypes, Inc., Provider Id 4321, provides free online access to genotyping assays from records in the Entrez database SNP. SNP records are selected using the SNP unique identifier. The URL to access assays at their site follows this pattern for each record: http://genotypinc.com/servlet/web.Gateway?source=NCBI.SNP&res=genotypAssay&ap1=rs[SNP ID]

```
-----Geotypes SNP links global info ---
prid:  4321
dbase: SNP

-----Begin Links ---
linkid: 1
uids:   7928656 2049045 1811350 1871598 7947824 681267 1947741
base:   http://genotypinc.com/servlet/web.Gateway?source=NCBI.SNP&res=genotypAssay&
rule:   ap1=rs&lo.id;
name:   Genotyping Assays
------
```

Example 3:

A record may be retrieved by more than one <Query>. When this happens, link assignment will be handled as described in Duplicate Links and Multiple Links.

If these queries are in different **linkids**, you can use **attr: preference** to indicate which link should be applied to the record. This is generally used in situations where the links for a subset of a range have a different URL pattern or different access restrictions. See Duplicate Links and Multiple Links for more information.

In the example below, the records included in LinkId 1 and LinkId 2 below will also be selected by LinkId 3.

The hypothetical LinkOut provider WebDatabase Co. provides links from Entrez Nucleotide to the *C. elegans* sequence database.

LinkId 1 describes links from particular Nucleotide records, specified by unique ID (GI) to a single record in the *C. elegans* database. The records are selected using **uids** and have a special **rule**. Because these records are also included in LinkId 3, **attr: preference** is used to indicate that only this link should be applied to these citations.

LinkId 2 describes links from all Nucleotide records on *C. elegans* published by J. Smith from 1997 to 1999 to a set of *C. elegans* records in PDF format. Because these records are also included in LinkId 3, **attr: preference** is used to indicate that only this link should be applied to these citations.

LinkId 3 provides links from all Nucleotide records on *C. elegans* to WebDatabase Co.'s *C. elegans* records, except for the records selected in LinkId 1 and LinkId 2.

Because both **LinkId 1** and **LinkId 2** describe specific requirements, they are listed before the general **LinkId 3**.

```
----- Nucleotide links ----
prid: 7777
dbase: nucleotide
!base: "http://www.webdatabase.com/cgi-bin/elegans?"

------
linkid: 1
uids:   3810674 1217583 1181594
base:   &base;
rule:   db=special&ID=A594E
attr:   preference

------
linkid: 2
query:  Caenorhabditis elegans [orgn] AND 1997:1999 [pdat] AND smith j [auth]
base:   &base;
rule:   auth_lookup=j-smith&view=pdf
attr:   full-text PDF
attr:   preference

------
linkid: 3
query:  Caenorhabditis elegans [orgn]
base:   &base;
rule:   an_lookup=&lo.pacc;&view=full
------
```

File Evaluation

After your application for inclusion in LinkOut has been accepted, prepare an identity file and sample resource files according to the instructions above. Resource files should contain links to at least five records from the selected database.

Validate the files using the LinkOut File Validation Utility. Email the files to linkout@ncbi.nlm.nih.gov.

Your files will be evaluated by the LinkOut team, and you will be contacted regarding any corrections. The evaluation process will continue until your files are substantially error free.

Account Assignment

When the submitted files are substantially error free, you will be assigned a ProviderId (PrId) and an approved name abbreviation (NameAbbr), and you will be given a password for an NCBI FTP account.

Please note that each provider will be given only one account at NCBI.

File Transfer

When you receive your account information, validate the files using the LinkOut File Validation Utility and transfer all files via FTP to the host FTP-private.ncbi.nih.gov. Place the files in the "holdings" directory of your FTP account. No subdirectories may be created in the holdings directory.

When files have been submitted, inform the LinkOut team at linkout@ncbi.nlm.nih.gov. Your files will be given a final evaluation before being placed in the production queue. From this point on, files will be processed automatically every day.

Links should appear in Entrez within 2 days of file submission. For instructions on viewing your links, see Using LinkOut. If 2 days have passed and you do not see your links in Entrez, please write to linkout@ncbi.nlm.nih.gov.

File Maintenance

Provider Responsibilities

Link providers are responsible for:

- maintaining their LinkOut files

- transferring any additions, changes or deletions of their links to NCBI

- updating files and informing NCBI when access rights are changed

- correcting broken or incorrect links in a timely manner

Providers may transfer new versions of current files or add new resource files at any time. It is the responsibility of the provider to keep files current and valid. Links are regenerated every day based on the resource files in each provider's directory. Therefore, providers must delete obsolete files from their holdings directory.

Additional provider responsibilities are described in LinkOut Policies: Provider Responsibilities.

Confirmation and Error Messages

Upon processing an updated file, NCBI will send an acknowledgment to the designated LinkOut contact. If you prefer not to receive this acknowledgment, please notify the LinkOut Team.

If files cannot be processed because of errors, a message with the subject line "LinkOut files uploaded to NCBI - Critical ERRORS!" will be sent to the LinkOut contact. In this case, please correct the files and resubmit them. If you have any questions about the errors, contact LinkOut at linkout@ncbi.nlm.nih.gov.

Provider Statistics

LinkOut collects statistics on the number of clicks on each providers's links in the LinkOut display.

Statistics can be emailed to the LinkOut contact monthly. If you would like to receive statistics, please notify the LinkOut Team.

Statistics may change for the first 2 weeks that they are available. After 2 weeks, statistics will be stable.

A sample statistics report is shown below.

```
The following is the usage statistics for your LinkOut links:
 (Hits = number of times users clicked your links)

Please note that statistics for the most recent month may change in the
next report.

Database: Gene

Total: 629

Month                    Number of Hits
-----                    --------------
8/2005                   126
7/2005                   195
6/2005                   142
5/2005                   155
4/2005                   11
```